

Throughput Maximization for Two-way Relay Channels with Energy Harvesting Nodes: The Impact of Relaying Strategies

Kaya Tutuncuoglu, Burak Varan, and Aylin Yener

Abstract—In this paper, we study the two-way relay channel with energy harvesting nodes. In particular, we find transmission policies that maximize the sum-throughput for two-way relay channels when the relay does not employ a data buffer. The relay can perform decode-and-forward, compress-and-forward, compute-and-forward or amplify-and-forward relaying. Furthermore, we consider throughput improvement by dynamically choosing relaying strategies, resulting in hybrid relaying strategies. We show that an iterative generalized directional water-filling algorithm solves the offline throughput maximization problem, with the achievable sum-rate from an individual or hybrid relaying scheme. In addition to the optimum offline policy, we obtain the optimum online policy via dynamic programming. We provide numerical results for each relaying scheme to support the analytic findings, pointing out to the advantage of adapting the instantaneous relaying strategy to the available harvested energy.

Index Terms—Energy harvesting nodes, two-way relay channel, decode/compute/compress/amplify-and-forward, hybrid relaying strategies, throughput maximization.

I. INTRODUCTION

Wireless networks consisting of energy harvesting nodes continue to gain significance in the area of green communications [2]–[26]. These networks *harvest* energy from external sources in an intermittent fashion, and consequently require careful management of the available energy.

There is considerable recent research on energy management for energy harvesting networks. Reference [4] considers an energy harvesting transmitter with energy and data arrivals, and an infinite size battery to store the harvested energy, and shows the optimality of a piecewise constant power policy for minimization of completion time of a file transfer. In [5], the throughput maximization problem is solved when the energy storage capacity of the battery is limited. It is shown that the transmission power policy is again piecewise constant, changing only when the battery is full or depleted. Extension of the model in [5] to fading channels is studied in [6] where a directional water-filling algorithm is shown to yield the optimum transmission policy. Reference [7] also considers throughput maximization for a fading channel under the same assumption. The impact of degradation and imperfections

of energy storage on the throughput maximizing policies is studied in [8]–[10]. The single user channel with an energy harvesting transmitter and an energy harvesting receiver is considered in [11], and decoding and sampling strategies for energy harvesting receivers is considered in [12]. Various multi-user energy harvesting networks have also been studied to date; including multiple access, broadcast, and interference channels with energy harvesting nodes [13]–[18]. In addition to these multi-user setups, variations of the energy harvesting relay channel are studied in [19]–[25], including multiple energy harvesting relays [26].

In this work, we study the simplest network setup that embodies a cooperative communication scenario with two-directional information flow, with the goal of identifying design insights unique to such scenarios. This leads to the investigation of bi-directional communication with energy harvesting nodes. Specifically, we study the so-called separated two-way relay channel¹ with energy harvesting nodes. The channel is separated in the sense that the users cannot hear each other directly, i.e., communication is only possible through the relay. This model is relevant and of interest for peer-to-peer communications, or for any scenario where a pair of nodes exchange information, and avails the relay node to implement strategies to convey both messages simultaneously. The two-way relay channel (TWRC) with conventional (non-energy-harvesting) nodes is studied with various relaying strategies such as amplify-and-forward, decode-and-forward, compress-and-forward [27]–[29], and compute-and-forward [30] in half-duplex [27], [28] and full-duplex [29], [31] models. It is observed that different relaying schemes outperform the others for different ranges of transmit powers.

In this paper, we identify transmission power policies for the energy harvesting two-way relay channel (EH-TWRC) which maximize the sum-throughput. The energy harvesting relay can perform amplify-and-forward, decode-and-forward, compress-and-forward, or compute-and-forward relaying. Due to intermittent energy availability, the channel calls for relaying strategies that adapt to varying transmit powers. For this purpose, we introduce a relay that can dynamically change its relaying strategy, resulting in what we term hybrid relaying strategies. We derive the properties of the optimal offline transmission policy, where energy arrivals are known non-causally, with the goal of gaining insights into its structure. Next, we show that an iterative generalized directional water-filling al-

This work was supported by NSF Grants CNS 09-64364 and CCF 14-22347. This work was presented in part at the International Conference on Communications, Workshop on Green Broadband Access: Energy Efficient Wireless and Wired Network Solutions, June 2013 [1].

The authors are with the department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: kaya@psu.edu; varan@psu.edu; yener@ee.psu.edu).

¹Usually referred to as the two-way relay channel, as we will in the sequel.

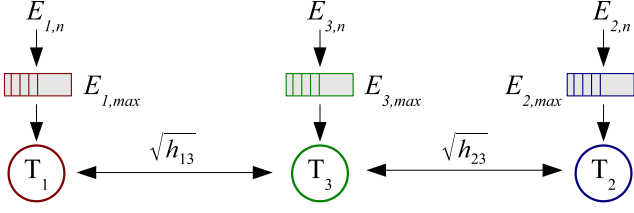


Fig. 1. The two-way relay channel with energy harvesting nodes (EH-TWRC).

gorithm solves the sum-throughput maximization problem for all relaying strategies. We next find the optimal online transmission policy by formulating and solving a dynamic program, where the energy states of the nodes are known causally. We compute optimal policies for different relaying strategies and provide numerical comparisons of their sum-throughputs. Our contribution includes generalization of directional water-filling [6] to an interactive communication scenario with multiple energy harvesting terminals in the offline setting, as well as the identification of optimal policies in the online setting. The interactive communication scenario considered in this paper is the catalyst that can drastically change the resulting power allocation algorithms in the energy harvesting setting. The two-way relay channel is the simplest multi terminal network model that demonstrates this interaction, and hence is the model considered. We observe that the relaying strategy has a significant impact on the optimum transmission policy, i.e., transmit powers and phase durations, and that hybrid relaying can provide a notable throughput improvement for the EH-TWRC.

The remainder of the paper is organized as follows. The system model is described in Section II. In Section III, a hybrid relaying scheme where the relay can alter its strategy depending on the instantaneous powers is introduced. In Section IV, the sum-throughput maximization problem is presented for an EH-TWRC, and is divided into subproblems that can be solved separately. In Section V, the iterative generalized directional water-filling algorithm is proposed to find an optimal policy for the EH-TWRC. The online policy based on dynamic programming is provided in Section VI. Numerical results are presented in Section VII. The paper is concluded in Section VIII.

II. SYSTEM MODEL

We consider an additive white Gaussian noise (AWGN) two-way relay channel with two source nodes, \$T_1\$ and \$T_2\$, that convey independent messages to each other through a relay node \$T_3\$. The two source nodes cannot hear each other directly, hence all messages are sent through the relay. The channels to and from a source node are reciprocal², with power gains \$h_{13}\$ between nodes \$T_1\$ and \$T_3\$ and \$h_{23}\$ between nodes \$T_2\$ and \$T_3\$. We consider the delay limited scenario, where the relay forwards messages as soon as they are received, and thus has no data buffer. The channel model is shown in Figure 1.

²While this assumption is for the sake of simplicity, we note that the results of this paper directly extend to models without reciprocity.

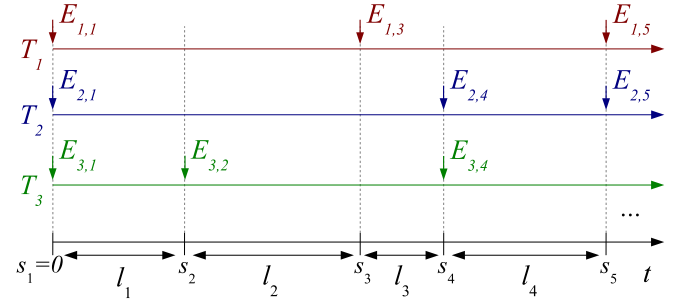


Fig. 2. The energy harvesting model for node \$T_j\$, \$j = 1, 2, 3\$.

All nodes \$T_1, T_2\$ and \$T_3\$ are powered by energy harvesting. Node \$T_j\$, \$j = 1, 2, 3\$, harvests \$E_{j,n} \geq 0\$ units of energy³ at time \$s_n\$, and stores it in a battery of energy storage capacity \$E_{j,max}\$. Any energy in excess of the storage capacity of the battery is lost. The initial charge of the batteries are represented with \$E_{j,1}\$, with \$s_1 = 0\$ by definition. The time between the \$n\$th and \$(n+1)\$th energy arrivals, referred to as the \$n\$th epoch in the sequel, is denoted by \$l_n = s_{n+1} - s_n\$. We remark that the model does not require all nodes to harvest energy packets simultaneously; but rather indicates that epochs are constructed as the intervals between any two energy arrivals. A node that is not receiving any energy at the \$n\$th harvest is set to have \$E_{j,n} = 0\$. The energy harvesting model is depicted in Figure 2.

We consider a transmission session of \$N\$ epochs, with length \$s_{N+1}\$, for which the energy harvesting profile consists of \$E_{j,n}\$ and \$s_n\$ for \$j = 1, 2, 3\$ and \$n = 1, \dots, N\$. In epoch \$n\$, \$n = 1, \dots, N\$, node \$T_j\$, \$j = 1, 2, 3\$, allocates an average power \$p_{j,n}\$ for transmission, i.e., a total energy of \$l_n p_{j,n}\$ is consumed for transmitting. Since the energy available to each node is limited by the energy harvested and stored in the respective battery, the energy harvesting profile determines the feasibility of \$p_{j,n}\$ for each node. Specifically, the transmission powers satisfy

$$p_{j,n} \leq \frac{B_{j,n}}{l_n}, \quad j = 1, 2, 3, \quad n = 1, \dots, N, \quad (1)$$

where \$B_{j,n}\$ is the energy available to node \$j\$ at the beginning of epoch \$n\$, which evolves as

$$B_{j,n+1} = \min\{E_{j,max}, B_{j,n} - l_n p_{j,n} + E_{j,n+1}\}. \quad (2)$$

In this work, similar to references [4]–[9], [11], [13]–[17], [19]–[26], the energy harvesting profile is known non-causally by all nodes, so that offline optimal policies and performance limits of the network can be found⁴. The communication overhead for conveying energy arrival information and power allocation decisions is considered to be negligible compared to the amount of data transferred in each epoch.

We consider the problem of finding the power policy which

³Recent efforts extend the discrete energy arrivals to continuous ones for the single user channel and concludes similar insights albeit with a more involved analysis [32].

⁴We provide the online policy with causal energy arrival information in Section VI.

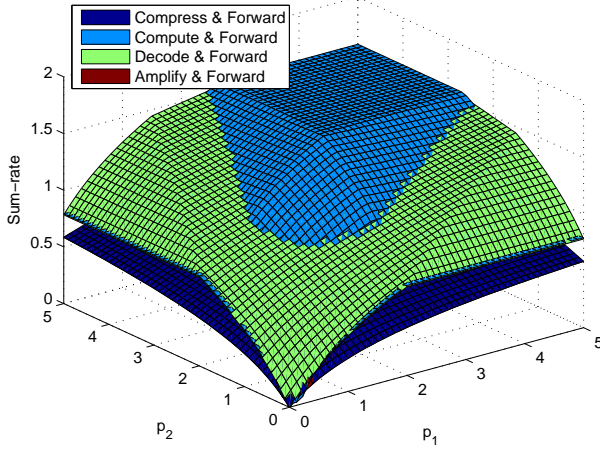


Fig. 3. Comparison of sum-rates for a symmetric full-duplex channel with $h_{13} = h_{23} = 1$ at $p_3 = 2$. Amplify-and-forward rates remain just below compress-and-forward and thus are not visible.

maximizes the sum-throughput of the system under different relaying strategies such as decode-and-forward, compute-and-forward, compress-and-forward and amplify-and-forward. In the next subsection, we present the rate regions for these relaying strategies.

A. Rate Regions with Average Power Constraints

We focus on a two-phase communication scheme, consisting of a multiple access phase from nodes T_1 and T_2 to T_3 and a broadcast phase from T_3 to T_1 and T_2 . This is referred to as multiple access broadcast (MABC) in [27], [28]. Its three phase counterpart, time division broadcast (TDBC) [27], [28], can be shown to perform no better than MABC in the absence of a direct channel between T_1 and T_2 , and is therefore omitted. For half-duplex nodes, the rates achievable with decode-and-forward, compress-and-forward, amplify-and-forward and compute-and-forward relaying schemes are derived in [27], [28]. These works consider nodes that are constrained by their instantaneous transmit powers, and do not consider total consumed energy, which depends on the duration of multiple access and broadcast phases. Since our model is energy-constrained, we revise the results of these work by scaling transmit powers with phase duration, thereby replacing instantaneous transmit powers with average transmit powers $p_{j,n}$. We denote the set of rate pairs achievable with average transmit powers p_1 , p_2 and p_3 and multiple access phase duration Δ as $\mathcal{R}_{HD}(p_1, p_2, p_3, \Delta)$ in the half-duplex case. The duration of the broadcast phase is $\bar{\Delta} = 1 - \Delta$. For full-duplex nodes, due to simultaneous multiple access and broadcast phases, there is no need for the time sharing factor Δ ; we use $\mathcal{R}_{FD}(p_1, p_2, p_3)$ to denote the achievable set of rate pairs. In this case, the full-duplex nodes can remove the self-interference term from the received signal, as in [29]. We use a subscript to denote the relaying strategy where necessary.

Decode-and-Forward: In this scheme, the relay decodes the messages of both source nodes in the multiple access phase, and transmits a function of the two messages in

the broadcast phase. Nodes T_1 and T_2 use the broadcast message along with their own messages to find the ones intended for them. For half-duplex nodes, the rate region $\mathcal{R}_{DF-HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n)$ in epoch n is defined by⁵

$$R_{1,n} \leq \min \left\{ \Delta_n C \left(\frac{h_{13} p_{1,n}}{\Delta_n} \right), \bar{\Delta}_n C \left(\frac{h_{23} p_{3,n}}{\bar{\Delta}_n} \right) \right\}, \quad (3a)$$

$$R_{2,n} \leq \min \left\{ \Delta_n C \left(\frac{h_{23} p_{2,n}}{\Delta_n} \right), \bar{\Delta}_n C \left(\frac{h_{13} p_{3,n}}{\bar{\Delta}_n} \right) \right\}, \quad (3b)$$

$$R_{1,n} + R_{2,n} \leq \Delta_n C \left(\frac{h_{13} p_{1,n} + h_{23} p_{2,n}}{\Delta_n} \right), \quad (3c)$$

where $C(p) = \frac{1}{2} \log(1 + p)$. With full-duplex radios, the two phases take place simultaneously, achieving instantaneous rates $(R_{1,n}, R_{2,n}) \in \mathcal{R}_{DF-FD}(p_{1,n}, p_{2,n}, p_{3,n})$ which are found by substituting $\Delta_n = \bar{\Delta}_n = 1$ in (3).

Compress-and-Forward: In this scheme, the relay transmits a compressed version of its received signal in the broadcast phase. The instantaneous rates $(R_{1,n}, R_{2,n}) \in \mathcal{R}_{CF-HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n)$, $0 \leq \Delta_n \leq 1$, for the MABC half-duplex case satisfy

$$R_{1,n} \leq \Delta_n C \left(\frac{(\sigma_y^{(1)})^2 h_{13} p_{1,n} / \Delta_n}{P_y^{(1)} (P_y^{(1)})^2 - (\sigma_y^{(1)})^2 (P_y^{(1)} - 1)} \right), \quad (4a)$$

$$R_{2,n} \leq \Delta_n C \left(\frac{(\sigma_y^{(1)})^2 h_{23} p_{2,n} / \Delta_n}{P_y^{(1)} (P_y^{(1)})^2 - (\sigma_y^{(1)})^2 (P_y^{(1)} - 1)} \right), \quad (4b)$$

for some $P_y^{(1)} \geq 0$ and $\sigma_y^{(1)} \geq 0$, where $P_y^{(1)} = h_{13} p_{1,n} / \Delta_n + h_{23} p_{2,n} / \Delta_n + 1$. The full-duplex rates $(R_{1,n}, R_{2,n}) \in \mathcal{R}_{CF-FD}(p_{1,n}, p_{2,n}, p_{3,n})$ are [29]

$$R_{1,n} \leq C \left(\frac{h_{13} p_{1,n}}{1 + \sigma_c^2} \right), \quad R_{2,n} \leq C \left(\frac{h_{23} p_{2,n}}{1 + \sigma_c^2} \right), \quad (5)$$

where $\sigma_c^2 = \max\{\sigma_{c1}^2, \sigma_{c2}^2\}$, and

$$\sigma_{c1}^2 = \frac{1 + h_{23} p_{2,n}}{2^2 R_{3,n}}, \quad \sigma_{c2}^2 = \frac{1 + h_{13} p_{1,n}}{2^2 R_{3,n}}, \quad (6a)$$

$$R_{3,n} = \min\{C(h_{13} p_{3,n}), C(h_{23} p_{3,n})\}. \quad (6b)$$

Amplify-and-Forward: In this scheme, the relay broadcasts a scaled version of its received signal. Since this is performed on a symbol-by-symbol basis, the time allocated for multiple access and broadcast phases are equal. The rate regions $\mathcal{R}_{AF}(p_{1,n}, p_{2,n}, p_{3,n})$ are found as

$$R_{1,n} \leq \Delta_n C \left(\frac{h_{13} h_{23} p_{1,n} p_{3,n}}{\Delta_n (h_{13} p_{1,n} + h_{23} (p_{2,n} + p_{3,n}) + \Delta_n)} \right), \quad (7a)$$

$$R_{2,n} \leq \Delta_n C \left(\frac{h_{13} h_{23} p_{2,n} p_{3,n}}{\Delta_n (h_{23} p_{2,n} + h_{13} (p_{1,n} + p_{3,n}) + \Delta_n)} \right), \quad (7b)$$

by substituting $\Delta_n = 0.5$ for the half-duplex case and $\Delta_n = 1$ for the full-duplex case.

⁵The power gains, transmit powers and harvested energy values are normalized in order to obtain an effective noise variance of 1 at each node. This is done by first scaling h_{13} and h_{23} to establish unit variance noise at nodes T_1 and T_2 , and subsequently scaling the transmit power, available energy and battery capacity at nodes T_1 and T_2 to yield a unit variance noise at T_3 .

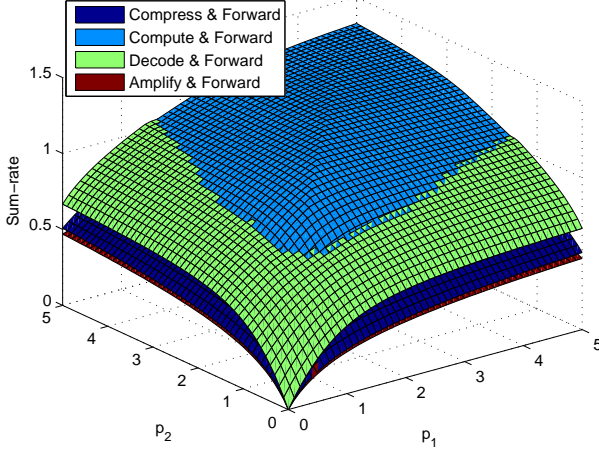


Fig. 4. Comparison of sum-rates for a symmetric half-duplex channel with $h_{13} = h_{23} = 1$ at $p_3 = 2$. Amplify-and-forward rates remain just below compress-and-forward and thus are barely visible.

Compute-and-Forward (Lattice Forwarding): In this scheme, nested lattice codes are used at the source nodes, and the relay decodes and broadcasts a function of the two messages received from the sources. Each source then calculates the intended message using the side information of its own [30]. The rate region $\mathcal{R}_{LF-HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n)$ achievable with this scheme for an MABC half-duplex relay consists of rates satisfying

$$R_{1,n} \leq \min \left\{ \frac{\Delta_n}{2} \log^+ \left(\frac{p_{1,n}}{p_{1,n} + p_{2,n}} + \frac{h_{13}p_{1,n}}{\Delta_n} \right), \bar{\Delta}_n C \left(\frac{h_{23}p_{3,n}}{\Delta_n} \right) \right\}, \quad (8a)$$

$$R_{2,n} \leq \min \left\{ \frac{\Delta_n}{2} \log^+ \left(\frac{p_{2,n}}{p_{1,n} + p_{2,n}} + \frac{h_{23}p_{2,n}}{\Delta_n} \right), \bar{\Delta}_n C \left(\frac{h_{13}p_{3,n}}{\Delta_n} \right) \right\}, \quad (8b)$$

where $\bar{\Delta}_n = 1 - \Delta_n$ and $\log^+(x) = \max\{\log x, 0\}$. The full-duplex rate region $\mathcal{R}_{LF-FD}(p_{1,n}, p_{2,n}, p_{3,n})$ can be evaluated by substituting $\Delta_n = \bar{\Delta}_n = 1$ in (8). In reference [30], it is shown that this strategy achieves within $\frac{1}{2}$ bits of TWRC capacity in each epoch.

It can be observed that the compute-and-forward rates are not jointly concave in transmit powers $p_{j,n}$, $j = 1, 2, 3$. This implies that time sharing between two sets of transmit powers $(p_{1,n}, p_{2,n}, p_{3,n})$ and $(\bar{p}_{1,n}, \bar{p}_{2,n}, \bar{p}_{3,n})$ with parameter λ , consuming average powers $\hat{p}_{j,n} = \lambda p_{j,n} + (1 - \lambda)\bar{p}_{j,n}$, $j = 1, 2, 3$, can yield rates $(R_{1,n}, R_{2,n}) \notin \mathcal{R}_{LF-FD}(\hat{p}_{1,n}, \hat{p}_{2,n}, \hat{p}_{3,n})$. To include rates achievable as such, we concavify the rate region by extending $\mathcal{R}_{LF-FD}(p_{1,n}, p_{2,n}, p_{3,n})$ to include all time-sharing combinations with average power $(p_{1,n}, p_{2,n}, p_{3,n})$, i.e.,

$$\mathcal{R}_{LF-FD}^C(p_{1,n}, p_{2,n}, p_{3,n}) = \left\{ (R_{1,n}, R_{2,n}) \left| R_{k,n} = \sum_i \lambda_i R_{k,n,i}, \right. \right.$$

$$\begin{aligned} (R_{1,n,i}, R_{2,n,i}) &\in \mathcal{R}_{LF-FD}(p_{1,n,i}, p_{2,n,i}, p_{3,n,i}), \\ \sum_i \lambda_i &= 1, \quad \sum_i \lambda_i p_{j,n,i} \leq p_{j,n}, \quad \lambda_i \geq 0, \\ j &= 1, 2, 3, \quad k = 1, 2 \end{aligned} \quad (9)$$

which we refer to as the concavified rate region. This extends to the half-duplex relaying region $\mathcal{R}_{LF-HD}(p_{1,n}, p_{2,n}, p_{3,n})$ by time sharing among Δ_n as well. With a slight abuse of notation, we will denote the concavified regions with $\mathcal{R}_{LF-FD}^C(p_{1,n}, p_{2,n}, p_{3,n})$ and $\mathcal{R}_{LF-HD}^C(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n)$ in the sequel. We note that all rates in the concavified region are achievable via time-sharing within an epoch, while the average powers within said epoch, and hence energy constraints, hold by definition. A formal proof of this concavification follows [16, Lem. 1] closely. In the sequel, we use the concavified region, though we do not reiterate the required time-sharing for clarity of exposition.

Since we are interested in maximizing sum-throughput, we compare the maximum achievable sum-rates for full- and half-duplex nodes employing the relaying strategies above in Figures 3 and 4, respectively. In these evaluations, a symmetric channel model normalized to yield $h_{13} = h_{23} = 1$, and a fixed relay power of $p_3 = 2$ is considered. It can be observed that different schemes may outperform based on the instantaneous transmit power, and thus the selection of the correct relaying scheme is of importance in an energy harvesting setting where transmit powers are likely to change throughout the transmission.

III. HYBRID SCHEMES

In Section II-A, it is observed that depending on the transmit powers, either one of the relaying strategies may yield the best instantaneous sum-rate. Due to the intrinsic variability of harvested energy, transmit powers may change significantly throughout the transmission period based on the energy availability of nodes. Consequently, a dynamic relay that chooses its relaying strategy based on instantaneous transmit powers of the nodes can potentially improve system throughput.

Another benefit of switching between relaying strategies is achieving time-sharing rates across strategies, e.g., switching between decode-and-forward and compute-and-forward strategies within an epoch, which can outperform both individual strategies with the same average power. An example of the benefits of time-sharing in a two-way relay channel is reference [33], where time-sharing between different operation modes is considered. In [33], a fixed relaying strategy is employed with different nodes transmitting at a time; while here we allow time-sharing between different relaying strategies.

The rates achievable with a hybrid strategy switching between the four relaying schemes in Figures 3 and 4 consist of the convex hull of the union of rate pairs achievable by the individual schemes. The rate region for the hybrid scheme is

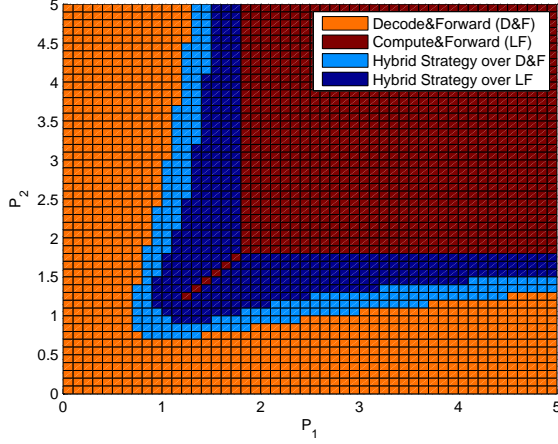


Fig. 5. Chosen relaying strategy for a symmetric half-duplex channel with $h_{13} = h_{23} = 1$ at $p_3 = 2$. The labels “over D&F” and “over LF” denote which of the two strategies is better by itself in that region.

expressed as

$$\mathcal{R}_{HYB}(p_1, p_2, p_3) = \left\{ (R_1, R_2) \left| R_k = \sum_i \lambda_i R_{k,i}, \right. \right. \\ \left. \sum_i \lambda_i = 1, \sum_i \lambda_i p_{j,i} \leq p_j, \lambda_i \geq 0, \right. \\ \left. (R_{1,i}, R_{2,i}) \in \mathcal{R}_{DF} \cup \mathcal{R}_{LF} \cup \mathcal{R}_{CF} \right. \\ \left. \cup \mathcal{R}_{AF}(p_{1,i}, p_{2,i}, p_{3,i}), j = 1, 2, 3, k = 1, 2 \right\}, \quad (10)$$

where \mathcal{R}_{DF} , \mathcal{R}_{LF} , \mathcal{R}_{CF} and \mathcal{R}_{AF} are the rate regions given in Section II-A with decode-and-forward, compute-and-forward, compress-and-forward and amplify-and-forward, respectively.

For the purpose of demonstration, we present the chosen relaying scheme that maximizes the instantaneous sum-rate for a half-duplex channel with fixed relay transmit power, $p_3 = 2$, in Figure 5. It can be observed that while decode-and-forward or compute-and-forward alone are chosen at the extremes, a time-sharing of the two strategies is favored in between. In this figure, the regions where the hybrid scheme uses time-sharing are shown in two shades of blue. We note that for these channel parameters, the remaining relaying schemes underperform these two for any choice of transmit powers.

With these observations, we conclude that policies with hybrid relaying strategies can instantaneously surpass the sum-rates resulting from individual relaying schemes for a considerable set of power vectors. Furthermore, time-sharing between relaying strategies may strictly outperform the best relaying strategy alone. Numerical results on the performance of optimal hybrid schemes in comparison with individual schemes are presented in Section VII.

IV. PROBLEM DEFINITION AND PROPERTIES OF THE OPTIMAL SOLUTION

We consider the problem of sum-throughput maximization for a session of N epochs. Since achievable rates are either

jointly concave in transmit powers or can be concavified by the use of time sharing as in (9), it follows that the optimal transmit powers remain constant within each epoch, as noted in [4, Lemma 2]. The power policy of the network consists of the power vectors $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$, where $\mathbf{p}_j = (p_{j,1}, p_{j,2}, \dots, p_{j,N})$, $j = 1, 2, 3$, and in the case of half-duplex relaying, the time sharing parameters Δ_n , $n = 1, \dots, N$. For the set of feasible power policies, we first present the following proposition, which is the multi-user extension of [5, Lemma 2]:

Proposition 1: There exists optimal average transmit powers $(\mathbf{p}_1^*, \mathbf{p}_2^*, \mathbf{p}_3^*)$ that do not yield a battery overflow at any of the nodes throughout the communication session.

Proof: Let $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ be a vector of transmit powers yielding battery overflows, i.e.,

$$\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} - E_{j,max} = E_{j,n}^{ovf} > 0 \quad (11)$$

for some j and n . For each battery overflow of amount $E_{j,n}^{ovf}$ at node T_j at the end of epoch n , let $\bar{p}_{j,n} = p_{j,n} + \frac{E_{j,n}^{ovf}}{l_n}$. For the remaining powers, let $\bar{p}_{j,n} = p_{j,n}$. The power policy defined by $(\bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2, \bar{\mathbf{p}}_3)$ does not overflow the battery at any time, and satisfies $\bar{p}_{j,n} \geq p_{j,n}$ for all j and n . Note that nodes consuming powers \bar{p}_j can achieve any rate pair that is achievable with less power, i.e.,

$$p_j \leq \bar{p}_j, j = 1, 2, 3 \\ \Rightarrow \mathcal{R}_{FD}(p_1, p_2, p_3) \subset \mathcal{R}_{FD}(\bar{p}_1, \bar{p}_2, \bar{p}_3), \quad (12a)$$

$$\Rightarrow \mathcal{R}_{HD}(p_1, p_2, p_3, \Delta) \subset \mathcal{R}_{HD}(\bar{p}_1, \bar{p}_2, \bar{p}_3, \Delta), \quad (12b)$$

for full-duplex and half-duplex nodes with $0 \leq \Delta \leq 1$, respectively. Therefore, the sum-rate obtained by $(\bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2, \bar{\mathbf{p}}_3)$ at any epoch n is no less than that of $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$. Hence, for any policy with battery overflows, we can find a policy performing at least as good without overflows. ■

We remark that even though (12) does not hold immediately, e.g., for the amplify-and-forward rates in (7), it holds by definition for the concavified rates in (9). By choosing $\lambda_1 = 1$ and $p_{j,n,1} < p_{j,n}$ in (9), a portion of the allocated power $p_{j,n}$ can equivalently be discarded at the node. Consequently, Proposition 1 applies to all concavified relaying schemes presented in Section II-A.

As a consequence of Proposition 1, we will restrict the feasible set of policies to those that do not overflow the battery without loss of generality. In epoch n , the nodes choose transmit powers $(p_{1,n}, p_{2,n}, p_{3,n})$, a time sharing parameter Δ_n , and a rate pair $(R_{1,n}, R_{2,n}) \in \mathcal{R}_{HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n)$ in the case of half-duplex radios. The objective is to maximize the sum-throughput of the TWRC within N epochs, where the transmit powers are constrained by harvested energy and the rates are constrained by the rate region. We express the EH-TWRC sum-throughput maximization problem

$$\max_{\mathbf{R}_1, \mathbf{R}_2, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \Delta} \sum_{i=1}^N l_i (R_{1,i} + R_{2,i}) \quad (13a)$$

$$\text{s.t. } (R_{1,n}, R_{2,n}) \in \mathcal{R}_{HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n), \quad (13b)$$

$$\sum_{i=1}^n l_i p_{j,i} - \sum_{i=1}^n E_{j,i} \leq 0, \quad (13c)$$

$$\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} \leq E_{j,max}, \quad (13d)$$

$$0 \leq \Delta_n \leq 1, \quad j = 1, 2, 3, \quad n = 1, 2, \dots, N, \quad (13e)$$

for half-duplex nodes, where $\mathbf{p}_j = (p_{j,1}, p_{j,2}, \dots, p_{j,N})$, $j = 1, 2, 3$, $\mathbf{R}_k = (R_{k,1}, R_{k,2}, \dots, R_{k,N})$, $k = 1, 2$, and $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_N)$. Here, (13d) is due to Proposition 1, and (13c) is equivalent to (1) given (13d). While the rates are a function of the powers of the nodes and the time sharing parameters $\Delta_n, n = 1, 2, \dots, N$, this dependency is now deferred to (13b), which is the constraint that ensures the rates are selected from the achievable region dictated by the power and time sharing parameters. The energy causality constraints given in (13c) ensure that the energy consumed by a node is not greater than the energy harvested up to that epoch. The no-overflow constraints given in (13d) ensure that the battery capacity is not exceeded. Any power policy $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ satisfying both (13c) and (13d) for all j and n is considered a feasible power policy. The problem for full-duplex nodes is attained by replacing (13b) with $(R_{1,n}, R_{2,n}) \in \mathcal{R}_{FD}(p_{1,n}, p_{2,n}, p_{3,n})$ and omitting the time-sharing variables $\Delta_n, n = 1, \dots, N$.

We next show that (13) can be decomposed by separating the maximization over $p_{1,n}, p_{2,n}$, and $p_{3,n}, n = 1, \dots, N$, and the maximization over $R_{1,n}, R_{2,n}, \Delta_n, n = 1, \dots, N$, as

$$\max_{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3} \max_{\mathbf{R}_1, \mathbf{R}_2, \Delta} \sum_{i=1}^N l_i (R_{1,i} + R_{2,i}) \quad (14a)$$

$$\text{s.t. } (R_{1,n}, R_{2,n}) \in \mathcal{R}_{HD}(p_{1,n}, p_{2,n}, p_{3,n}, \Delta_n), \quad (14b)$$

$$\sum_{i=1}^n l_i p_{j,i} - \sum_{i=1}^n E_{j,i} \leq 0, \quad (14c)$$

$$\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} \leq E_{j,max}, \quad (14d)$$

$$0 \leq \Delta_n \leq 1, \quad j = 1, 2, 3, \quad n = 1, 2, \dots, N. \quad (14e)$$

Note that only the constraints in (14b) pertain to the parameters of the second maximization. Next, we observe that the constraints in (14b) are separable in n , and the objective is a linear function of $R_{1,n}$ and $R_{2,n}$. Hence, the second maximization can be carried out separately for each n , i.e., in an epoch-by-epoch fashion, yielding the separated problem

$$\max_{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3} \sum_{i=1}^N l_i R_s(p_{1,i}, p_{2,i}, p_{3,i}) \quad (15a)$$

$$\text{s.t. } \sum_{i=1}^n l_i p_{j,i} - \sum_{i=1}^n E_{j,i} \leq 0, \quad (15b)$$

$$\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} \leq E_{j,max}, \quad (15c)$$

$$j = 1, 2, 3, \quad n = 1, 2, \dots, N, \quad (15d)$$

where $R_s(p_{1,i}, p_{2,i}, p_{3,i})$ is the solution to

$$\max_{R_{1,i}, R_{2,i}, \Delta_i} R_{1,i} + R_{2,i} \quad (16a)$$

$$\text{s.t. } (R_{1,i}, R_{2,i}) \in \mathcal{R}_{HD}(p_{1,i}, p_{2,i}, p_{3,i}, \Delta_i), \quad (16b)$$

$$0 \leq \Delta_i \leq 1, \quad (16c)$$

within a single epoch i with fixed powers $(p_{1,i}, p_{2,i}, p_{3,i})$. This implies that the optimal transmit rates within each epoch are the sum-rate maximizing rates for the given transmit powers within that epoch. Thus, we refer to the function $R_s(p_{1,i}, p_{2,i}, p_{3,i})$ as the *maximum epoch sum-rate*. For full-duplex nodes, the maximum epoch sum-rate is found by solving

$$\max_{R_{1,i}, R_{2,i}} R_{1,i} + R_{2,i} \quad (17a)$$

$$\text{s.t. } (R_{1,i}, R_{2,i}) \in \mathcal{R}_{FD}(p_{1,i}, p_{2,i}, p_{3,i}) \quad (17b)$$

instead, and the power policy optimization is identical to (15). We next show a property of policies that solve the problem in (13).

Lemma 1: There exists an optimal policy which depletes the batteries of all nodes at the end of transmission.

Proof: Let $(\mathbf{R}_1, \mathbf{R}_2, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ be a transmission policy which leaves energy \mathcal{E}_j in the battery of node j at the end of transmission. Consider the transmission policy $(\bar{\mathbf{R}}_1, \bar{\mathbf{R}}_2, \bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2, \bar{\mathbf{p}}_3)$ which has $\bar{p}_{j,N} = p_{j,N} + \mathcal{E}_j/l_N$, and equals the original policy elsewhere. Hence, this policy expends the remaining energy in the battery of T_j in the last epoch, depleting the batteries. We have $\bar{R}_{k,n} = R_{k,n}$ for $n = 1, 2, \dots, N-1$ and $\bar{R}_{k,N} \geq R_{k,N}$, for $k = 1, 2$, due to (12). Therefore, the sum-throughput of the new policy cannot be lower than that of the original policy. ■

V. IDENTIFYING THE OPTIMAL POLICY

Now that we have formulated the problem and identified some necessary properties of the optimal policy, we next find the optimal power policy for the EH-TWRC. We establish this using a generalization of the directional water-filling algorithm in [6], which gives the optimal policy for a single transmitter fading channel. In this section, we show the optimality of the generalized directional water-filling algorithm and verify its convergence.

A. Solution of the EH-TWRC Sum-Throughput Maximization Problem

To find the optimal policy, we first find the maximum epoch sum-rate by solving (16) and (17) for half-duplex and full-duplex nodes, respectively. The following property of $R_s(p_{1,i}, p_{2,i}, p_{3,i})$ can be immediately observed for any relaying scheme.

Lemma 2: The maximum epoch sum-rate $R_s(p_{1,i}, p_{2,i}, p_{3,i})$ is jointly concave in transmit powers $p_{1,i}$, $p_{2,i}$, and $p_{3,i}$.

Proof: Proof follows from the concavity of objectives (16a) and (17a), and the convexity of constraint sets (16b) and (17b). Let (R_1, R_2) and (\bar{R}_1, \bar{R}_2) denote two feasible rate pairs, and Δ and $\bar{\Delta}$ their time-sharing parameters for transmit

powers (p_1, p_2, p_3) and $(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3)$, respectively. Let $\bar{R}_k = \alpha R_k + (1 - \alpha)\bar{R}_k$, $k = 1, 2$, $\bar{p}_j = \alpha p_j + (1 - \alpha)\tilde{p}_j$, and $\bar{\Delta} = \alpha\Delta + (1 - \alpha)\tilde{\Delta}$, $j = 1, 2, 3$ denote the convex combination of the policies with parameter $0 \leq \alpha \leq 1$. Then, for all relaying schemes, $(\bar{R}_1, \bar{R}_2) \in \mathcal{R}_{FD}(\bar{p}_1, \bar{p}_2, \bar{p}_3)$ or $(\bar{R}_1, \bar{R}_2) \in \mathcal{R}_{HD}(\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{\Delta})$ follows either from the definition of the rate region, or from (9). ■

As a consequence of Lemma 2, (15) is a convex program. We next provide the *iterative generalized directional water-filling algorithm* to compute the optimal power policy. Consider the power allocation problem in (15) for an arbitrary relaying scheme with the maximum epoch sum-rate $R_s(p_1, p_2, p_3)$. Here, the constraints in (15b) and (15c) are separable among $j = 1, 2, 3$. Hence, a block coordinate descent algorithm, i.e., alternating maximization, can be employed [34]. In each iteration, the power allocation problem for node T_j , $j = 1, 2, 3$, given by

$$\max_{\mathbf{p}_j \geq 0} \sum_{n=1}^N l_n R_s(p_{1,n}, p_{2,n}, p_{3,n}) \quad (18a)$$

$$\text{s.t.} \quad \sum_{i=1}^n l_i p_{j,i} - \sum_{i=1}^n E_{j,i} \leq 0, \quad (18b)$$

$$\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} \leq E_{j,max}, \quad (18c)$$

$$n = 1, 2, \dots, N, \quad (18d)$$

is solved while keeping the remaining power levels \mathbf{p}_k , $k \neq j$, constant. This is a convex single user problem, and the solution satisfies the KKT stationarity conditions and complementary slackness conditions [34]

$$l_n \frac{\partial R_s(p_{1,n}, p_{2,n}, p_{3,n})}{\partial p_{j,n}} - l_n \sum_{i=n}^N (\lambda_i - \beta_i) + \gamma_n = 0, \quad (19a)$$

$$\lambda_n \left(\sum_{i=1}^n l_i p_{j,i} - \sum_{i=1}^n E_{j,i} \right) = 0, \quad (19b)$$

$$\beta_n \left(\sum_{i=1}^n E_{j,i} - \sum_{i=1}^{n-1} l_i p_{j,i} - E_{j,max} \right) = 0, \quad \gamma_n p_{j,n} = 0, \quad (19c)$$

for all $n = 1, \dots, N$ where $\lambda_n \geq 0$, $\beta_n \geq 0$ and $\gamma_n \geq 0$ are the Lagrange multipliers for energy causality, battery capacity and transmit power non-negativity constraints, respectively. Hence, the optimal transmit power policy for T_j , i.e., \mathbf{p}_j , is the solution to

$$\frac{\partial R_s(p_{1,n}, p_{2,n}, p_{3,n})}{\partial p_{j,n}} = \sum_{i=n}^N (\lambda_i - \beta_i) - \frac{\gamma_n}{l_n} \quad (20)$$

for all $n = 1, \dots, N$ which follows from (19a). Note that due to (19b) and (19c), the Lagrange multipliers are nonzero only when the respective constraints are met with equality.

We argue that the solution to (20) can be interpreted as a generalization of the directional water-filling algorithm [6] similar to the case in [16]. In [6], optimal transmit powers are found by treating the available energy in each epoch as water, and letting water levels equalize by flowing in the forward direction only. The associated algorithm is termed directional

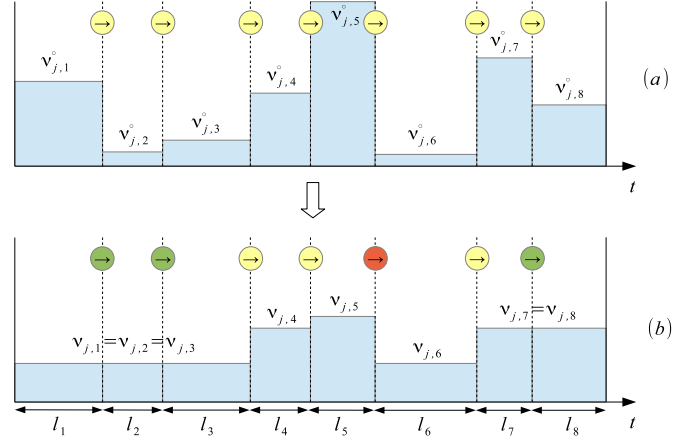


Fig. 6. Depiction of generalized directional water-filling for T_j with $N = 8$ epochs. Note that the battery of the node is full at the end of the 5th epoch, preventing further energy flow into the 6th epoch.

water-filling. Here, we instead define the *generalized water levels* for T_j as

$$\nu_{j,n}(p_{j,n}) = \left(\frac{\partial R_s(p_{1,n}, p_{2,n}, p_{3,n})}{\partial p_{j,n}} \right)^{-1}. \quad (21)$$

The following properties of $\nu_{j,n}$ are readily observed for the optimal policy: (a) while $p_{j,n} > 0$, the water levels remain constant among epochs unless the battery is empty or full, increasing only when the battery is empty, and decreasing only when the battery is full, and (b) if a positive solution to

$$\frac{\partial R_s(p_{1,n}, p_{2,n}, p_{3,n})}{\partial p_{j,n}} = \sum_{i=n}^N (\lambda_i - \beta_i) \quad (22)$$

does not exist, then $p_{j,n} = 0$ and $\gamma_n \geq 0$. These properties imply that the optimal policy can be found by performing directional water-filling using the generalized water levels in (21), and calculating the corresponding transmit powers $p_{j,n}$. Water flow is only in the forward direction and the corresponding energy flow is bounded by $E_{j,max}$ for node T_j . Hence, the flow between two neighboring epochs stops when water levels in (21) are equalized or when the total energy flow reaches $E_{j,max}$. The initial water levels are found by substituting the initial transmit powers $p_{j,n}^o = E_{j,n}/l_n$ in (21). This algorithm yields transmit powers that satisfy the two properties above by construction. An example of generalized directional water-filling is depicted in Figure 6.

The *iterative generalized directional water-filling* (IGDWF) algorithm employs generalized directional water-filling sequentially for each user until all power levels \mathbf{p}_j , $j = 1, 2, 3$, converge, i.e., alternating maximization. Although optimization is carried on separately for a single user at each iteration, the transmit powers of all users interact through the generalized water levels in (21). Starting from the initial values $p_{j,n}^{(0)} = E_{j,n}/l_n$, the k th iteration of the algorithm, optimizing $\mathbf{p}_j^{(k)}$ for $j = (k \bmod 3) + 1$, is given in Algorithm 1.

Remark 1: At each iteration of the IGDWF algorithm, the water flow out of each of the N epochs can be found using a

Algorithm 1 Iteration k of Iterative Generalized Directional Water-filling

- 1) Let $j = (k \bmod 3) + 1$, $p_{j,n}^{(k)} = p_{j,n}^{(0)}$, $p_{\ell,n}^{(k)} = p_{\ell,n}^{(k-1)}$ for $\ell \neq j$, $\delta_n = E_{j,n}$, $n = 1, \dots, N$.
 - 2) **for** $n = 2, \dots, N$, **do**
 - Find the set $\mathcal{E} = \{E_{\Delta} \geq 0 \mid \nu_{j,n-1}(p_{j,n-1}^{(k)} - \frac{E_{\Delta}}{l_n}) = \nu_{j,n}(p_{j,n}^{(k)} + \frac{E_{\Delta}}{l_n}), \delta_n + E_{\Delta} \leq E_{j,max}\}$,
 - if** $\mathcal{E} = \emptyset$ and $\nu_{j,n-1}(p_{j,n-1}^{(k)}) > \nu_{j,n}(p_{j,n}^{(k)})$, **then**
 - assign $\mathcal{E} = \{E_{j,max} - \delta_n\}$,
 - Find $E_{\Delta} \in \mathcal{E}$ and assign $p_{j,n-1}^{(k)} = p_{j,n-1}^{(k)} - \frac{E_{\Delta}}{l_{n-1}}$, $p_{j,n}^{(k)} = p_{j,n}^{(k)} + \frac{E_{\Delta}}{l_n}$, $\delta_n = \delta_n + E_{\Delta}$
 - such that $\|\mathbf{p}_j^{(k)} - \mathbf{p}_j^{(k-1)}\|$ is minimized.
 - end for**
 - 3) Repeat 2 until $\nu_{j,n-1}(p_{j,n-1}^{(k)}) \leq \nu_{j,n}(p_{j,n}^{(k)})$ or $\delta_n = E_{j,max}$ for all n .
-

binary search. This requires updating at most N water levels following each epoch. Hence, the computational complexity of each iteration is $O(N^2)$, i.e., quadratic in the number of epochs.

B. Convergence of the IGDWF Algorithm

For the alternating maximization in Section V-A to converge to an optimal policy, it is sufficient that the feasible set is the intersection of convex constraints that are separable among $j = 1, 2, 3$, and the continuously differentiable objective yields a unique maximum in each iteration [34, Prop. 2.7.1]. In this case, the objective in (15a) is concave and continuously differentiable for all relaying strategies, with compute-and-forward satisfying this condition after the concavification in (9). The feasible set (15b)-(15c) is separable among $j = 1, 2, 3$ as well. However, the objective does not necessarily yield a unique maximum at each iteration since it is not strictly concave in transmit powers. To overcome this, we introduce the unconstrained variables $\mathbf{s}_j = (s_{j,1}, \dots, s_{j,N})$ for $j = 1, 2, 3$, and modify the objective in (18a) as

$$f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3) = \sum_{n=1}^N l_n R_s(p_{1,n}, p_{2,n}, p_{3,n}) - \epsilon_1 \|\mathbf{p}_1 - \mathbf{s}_1\|^2 - \epsilon_2 \|\mathbf{p}_2 - \mathbf{s}_2\|^2 - \epsilon_3 \|\mathbf{p}_3 - \mathbf{s}_3\|^2, \quad (23)$$

where $\epsilon_j > 0$, $j = 1, 2, 3$ are arbitrarily small parameters. The objective in (23) is maximized by a unique \mathbf{p}_j in each iteration with $j = 1, 2$, or 3 . The iterations optimizing \mathbf{s}_j trivially yield the unique solution $\mathbf{s}_j = \mathbf{p}_j$. Therefore, the problem now satisfies the convergence property for alternating maximization, and converges to the global maximum of (15) [34, Ex. 2.7.2].

Note that through (23) and the arbitrarily small ϵ_j , we essentially introduce *resistance* to the iterative algorithm. That is, if the original objective in (15a) yields multiple solutions for some j , the objective in (23) has a unique solution that is closest to the previous value of \mathbf{p}_j . Consequently, if there exists more than one optimal solution to (18) at one of the iterations for some j , the power policy \mathbf{p}_j that is closest to

the previous one is chosen. This is ensured by choosing the flow amount E_{Δ} which minimizes $\|\mathbf{p}_j^{(k)} - \mathbf{p}_j^{(k-1)}\|$ in Step 2 of Algorithm 1.

VI. ONLINE POWER POLICY WITH DYNAMIC PROGRAMMING

The power allocation policy we have considered so far is an offline policy, in the sense that the energy harvest amounts and times are known to all nodes in advance. Although the offline approach is useful for predictable energy harvesting scenarios [35] and as a benchmark, it is also meaningful to develop policies that only rely on past and current energy states, i.e., causal information only. We refer to such transmission policies as *online* policies. Recent efforts that consider online algorithms for energy harvesting nodes in various channel models include [6], [7], [18], [23], [36]–[38]. Building upon the previous work, in this section, we find the optimal online policy for power allocation in the two-way relay channel.

The epoch length l_n indicates that no energy will be harvested for a duration of l_n after the n th energy arrival. Therefore, in the online problem, the epoch lengths are not known by the nodes causally. Instead, we divide the transmission period into time slots of length τ , and recalculate transmit powers at the beginning of each time slot. We assume that each energy harvest takes place at the beginning of some time slot. Note that with smaller τ , this model gets arbitrarily close to the general model in Section II. We assume that harvests $E_{j,n}$ in time slot n are independent and identically distributed. In time slot n , nodes T_1 , T_2 and T_3 have access to previous energy harvests $E_{j,i}$, $j = 1, 2, 3$, and $i = 1, \dots, n$. The nodes decide on transmit powers $p_{j,n}$, $j = 1, 2, 3$, through *actions*

$$p_{j,n} = \phi_{j,n}(\{E_{k,i}; k = 1, 2, 3, i = 1, \dots, n\}), \quad (24)$$

where $\{E_{k,i}; k = 1, 2, 3, i = 1, \dots, n\}$ denotes all energy arrivals prior to, and including, time slot n . Each time slot with transmit powers $\{p_{j,n}\}$ contribute to the additive objective through the sum-rate function $R_s(p_{1,n}, p_{2,n}, p_{3,n})$ in (16) and (17) for full-duplex and half-duplex modes, respectively. We consider the problem of finding the optimal set of actions for this setting, which can be formulated as the following dynamic program [39]:

$$V(\{E_{j,i}\}_{i=1}^n) = \max_{\phi_{1,n}, \phi_{2,n}, \phi_{3,n}} R_s(\phi_{j,n}(\{E_{k,i}\})) + \mathbb{E} \left[\sum_{i=n+1}^N V(\{E_{j,m}\}_{m=1}^i) \right]. \quad (25)$$

Here N is the number of time slots and $\mathbb{E}[\cdot]$ denotes the conditional expectation over remaining energy harvests $\{E_{j,i}\}_{i=n+1}^N$ given the previous harvests as $\{E_{j,i}\}_{i=1}^n$.

Note that the dynamic program outlined by (25) is computationally difficult due to the dimension of the problem. However, for the case of i.i.d. energy harvests that we consider, it can be simplified by restricting to actions that only utilize current battery state. This is due to the expectation in

(25) being independent of past energy harvests. This implies simplifying the actions in (24) as

$$p_{j,n} = \phi_{j,n}(E_{1,n}^{bat}, E_{2,n}^{bat}, E_{3,n}^{bat}), \quad (26a)$$

$$E_{j,n}^{bat} = \sum_{i=1}^{n-1} (E_{j,i} - \tau p_{j,i}) + E_{j,n}, \quad (26b)$$

where $E_{j,n}^{bat}$ is the battery state of T_j at the beginning of slot n . The solution to (25) and (26) provides the optimal online power policy for finite horizon, which we compare with the offline policy in Section VII.

To further simplify the problem, we additionally consider the infinite horizon problem where the optimal actions are time-invariant. We formulate this as a discounted dynamic program with the Bellman equation

$$V(\{E_j^{bat}\}) = \max_{\phi_1, \phi_2, \phi_3} R_s(\phi_j(\{E_k^{bat}\})) + \beta \mathbb{E}[V(\{E_j^{bat} - \tau \phi_j(\{E_k^{bat}\}) + E_j\})], \quad (27)$$

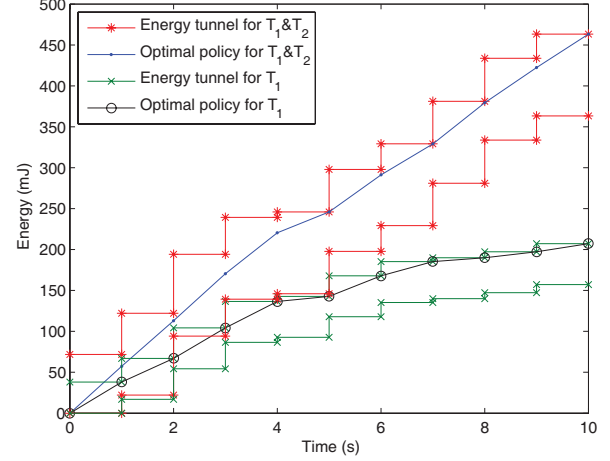
where $\phi_j = \phi_{j,n}$ for all n and the expectation is over E_j , $j = 1, 2, 3$. This equation can be solved with value iteration [39]. Namely, starting from arbitrary initial actions, all actions ϕ_j are updated as the arguments that maximize (27), and value functions $V(\{E_j^{bat}\})$ are updated as in (27), until all actions converge to some ϕ_j^* . Here, the discount factor $\beta < 1$ ensures that the values $V(\{E_j^{bat}\})$ remain bounded [39]. The resulting actions yield an online policy that is optimal under the action restrictions in (26) and infinite transmission assumption. Hence, we refer to this policy as the optimal online policy for an infinite horizon.

Remark 2: Each value iteration step requires K^3 value updates, where K is the number of transmit power values after discretization. Hence, the running time of the finite horizon algorithm is $O(NK^3)$ for N epochs, and storing the optimal policy requires $O(NK^3)$ space. On the other hand, the optimal policy is time-invariant for the infinite horizon case, and requires only $O(K^3)$ space.

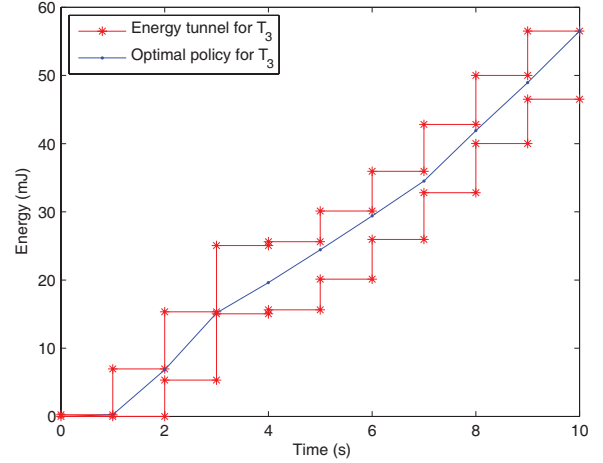
VII. NUMERICAL RESULTS

In this section, we demonstrate the optimal policies for the two-way relay channel and compare the performance of the schemes in Section II-A and Section III in the EH setting. In simulations, energy arrivals to node T_j are generated independently from a uniform distribution over $[0, E_{h,j}]$ for $j = 1, 2, 3$, with unit epoch lengths $l_n = 1$ s. The noise density is 10^{-19} W/Hz at all nodes and the bandwidth is 1 MHz.

Examples for the optimal transmit power policies found using the algorithm described in Section V are shown in Figures 7–10 for decode-and-forward relaying. In each figure, cumulative energy consumed by the nodes for transmission are plotted, the derivative of which yields the average transmit powers of the nodes in each epoch. In the figures, $T_1 \& T_2$ stands for the total cumulative energy of the nodes T_1 and T_2 , and MAC fraction represents the fraction of the multiple access phase, i.e., Δ_n . We remark that concavified sum-rate functions are used for the simulations, and average transmit powers are shown in the plots for clarity. Pairs of staircases, shown in red and green, represent energy causality and battery



(a)

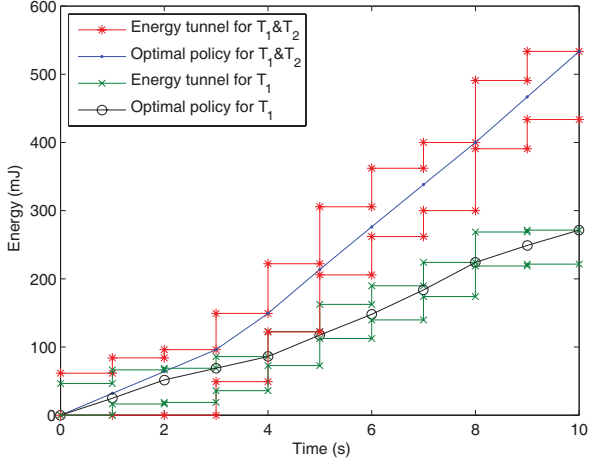


(b)

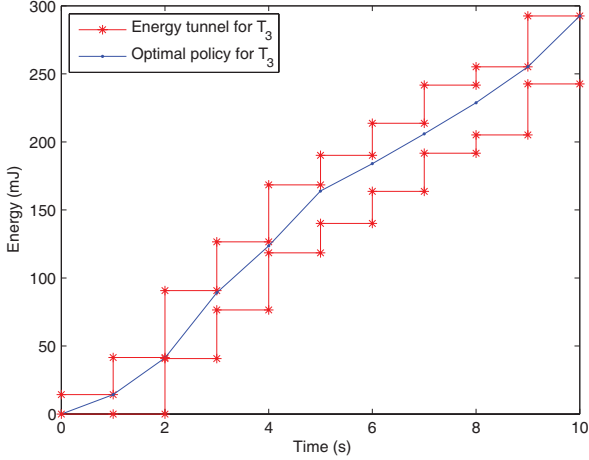
Fig. 7. Optimal cumulative harvested energy and consumed energy policies for (a) node T_1 and sum of T_1 and T_2 , and (b) node T_3 , for an asymmetric full-duplex channel with decode-and-forward relaying, $h_{13} = -110$ dB, $h_{23} = -116$ dB, peak energy harvesting rates $E_{h,1} = E_{h,2} = 50$ mJ and $E_{h,3} = 10$ mJ, battery sizes $E_{1,max} = E_{2,max} = 50$ mJ and $E_{3,max} = 10$ mJ.

capacity constraints on the cumulative power, which is referred to as the feasible energy tunnel [5]. A feasible policy remains between these two constraints throughout the transmission period. Figures 7 and 8 are plotted for full-duplex nodes while Figures 9 and 10 are plotted for half-duplex nodes. Both scenarios are considered for an asymmetric EH-TWRC with $h_{13} \neq h_{23}$ in Figures 7 and 9, and for a symmetric EH-TWRC with $h_{13} = h_{23}$ in Figures 8 and 10.

We remark that unlike previous work with simpler channel models, e.g., [4], [5], [13], [15], the optimal cumulative energy or sum-power policy is not necessarily the shortest path that traverses the feasible tunnel. Figure 7 shows a setting with $E_{h,1} = E_{h,2} = 50$ mJ and $E_{h,3} = 10$ mJ, i.e., the relay is energy deprived compared to T_1 and T_2 . Hence, energy efficiency is critical for the relay while this is not necessarily the case for the remaining nodes that are relatively energy-rich.



(a)

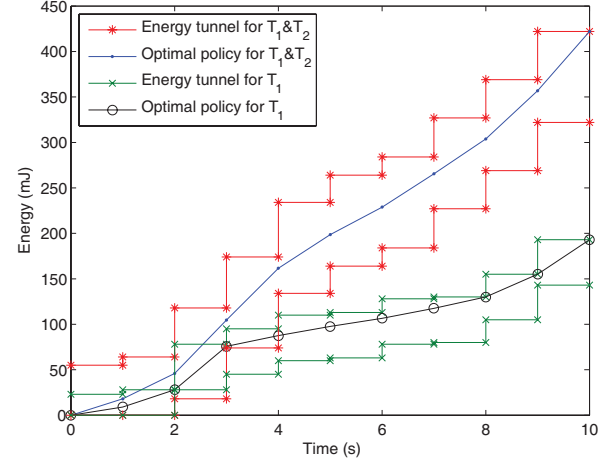


(b)

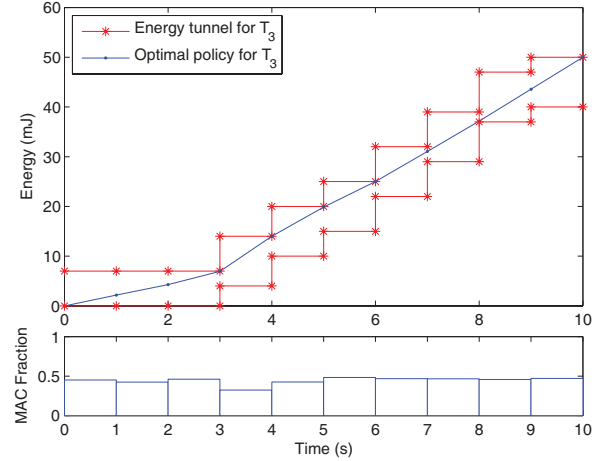
Fig. 8. Optimal cumulative harvested energy and consumed energy policies for (a) node T_1 and sum of T_1 and T_2 , and (b) node T_3 , for a symmetric full-duplex channel with decode-and-forward relaying and $h_{13} = h_{23} = -110$ dB, peak energy harvesting rates $E_{h,1} = E_{h,2} = E_{h,3} = 50$ mJ, battery sizes $E_{1,max} = E_{2,max} = E_{3,max} = 50$ mJ.

This results in the optimal policy being largely dictated by the relay. Note that in Figure 7, the relay follows a cumulative energy that resembles the shortest path through the feasible energy tunnel, while for T_1 and T_2 this is not the case. In contrast, in Figure 8, the multiple access phase is more likely to be limiting because the sum-rate with equal transmit powers at all nodes is limited by the sum-rate constraint of the multiple access phase, see (3c). Thus, the total cumulative energy, denoted with $T_1 \& T_2$ in Figure 8(a), follows the shortest path within the tunnel, similar to the optimal policy for the multiple access channel in [13]. However, broadcast powers do not yield binding constraints, implying that contrary to the energy harvesting models previously studied, e.g., [4], [5], the optimal policy for the EH-TWRC is not necessarily unique.

Comparing Figures 9 and 10, which show optimal policies for the half-duplex model, we observe that the time division parameters Δ_n play an important role in helping energy deprived nodes. By properly selecting Δ_n , the effect



(a)

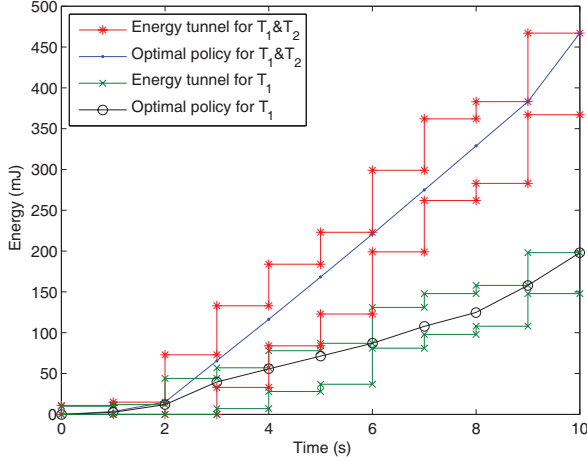


(b)

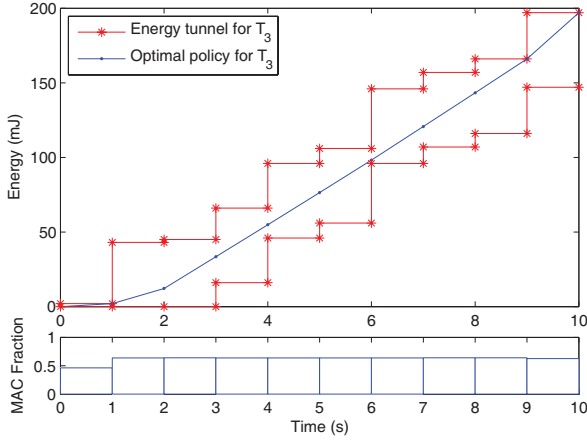
Fig. 9. Optimal cumulative harvested energy and consumed energy policies for (a) node T_1 and sum of T_1 and T_2 , and (b) node T_3 , for an asymmetric half-duplex channel with decode-and-forward relaying and $h_{13} = -110$ dB, $h_{23} = -116$ dB, peak energy harvesting rates $E_{h,1} = E_{h,2} = 50$ mJ, $E_{h,3} = 10$ mJ and battery sizes $E_{1,max} = E_{2,max} = 50$ mJ and $E_{3,max} = 10$ mJ.

of unbalanced energy harvests at the sources and the relay can be mitigated. However, this still does not imply the shortest path is optimal for each node. This is due to the interplay of transmit powers though the joint rate function in the objective. Hence, whenever the transmit power changes for one user due to a full battery or an empty battery, the transmit powers of other users are affected as well. Examples to this phenomena can be found in Figure 9, at $t = 3, 4$ s, where the energy depletion in T_3 is observed to affect the transmit powers of T_1 and T_2 , and in Figure 10, at $t = 2$ s, where the energy depletion in T_1 and T_2 is observed to affect the transmit power of T_3 .

Remark 3: Similar results were observed for compress-and-forward, compute-and-forward, and amplify-and-forward relaying through simulations. We observed that identical energy harvesting profiles and channel parameters yield transmit pow-



(a)



(b)

Fig. 10. Optimal cumulative harvested energy and consumed energy policies for (a) node T_1 and sum of T_1 and T_2 , and (b) node T_3 , for a symmetric half-duplex channel with decode-and-forward relaying and $h_{13} = h_{23} = -110$ dB, peak energy harvesting rates $E_{h,1} = E_{h,2} = E_{h,3} = 50$ mJ, battery sizes $E_{1,max} = E_{2,max} = E_{3,max} = 50$ mJ.

ers that only differ slightly among relaying schemes. However, the multiple access phase fractions, $\Delta_n, n = 1, 2, \dots, N$, differ notably among relaying schemes in order to achieve matching multiple access and broadcast rates within each epoch. Due to the similarity of the transmit power policies, to avoid repetition, we omit the plots for these schemes.

Next, we compare the performance of the optimal offline and online policies with upper and lower bounds for a decode-and-forward relay. We obtain a non-energy-harvesting upper bound by providing the total energy harvested by each node at the beginning of the transmission without a battery restriction. We also present two naïve transmit power policies, namely the hasty policy and the constant power policy, as lower bounds. The former policy, also referred to as the spend-as-you-get algorithm [40], consumes all harvested energy immediately within the same epoch. The latter policy chooses the average harvest rate at each node as the desired transmit

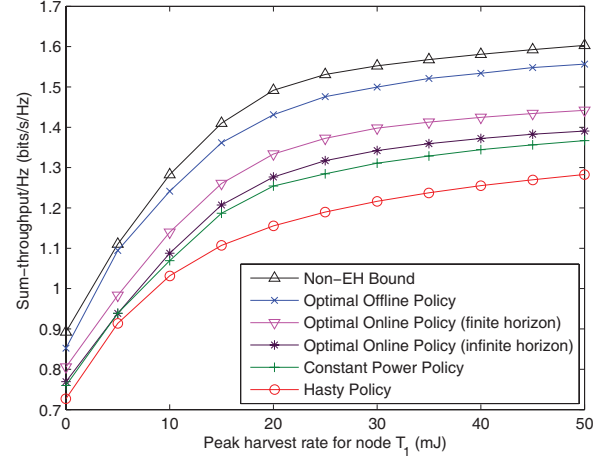


Fig. 11. Sum-throughput with optimal power allocations for decode-and-forward relaying compared with a non-EH upper bound, hasty policy and constant power policy.

power, and transmits with this power whenever possible. For both naïve policies, the phase fraction parameters Δ_n that maximize the instantaneous sum-rate for the given transmit powers are chosen within each epoch. We consider a half-duplex EH-TWRC with $h_{13} = h_{23} = -110$ dB, and choose the energy harvests for node T_j to be independent and uniformly distributed over $[0, E_{h,j}]$ where $E_{h,2} = 50$ mJ and $E_{h,3} = 20$ mJ are the peak harvest rates. The infinite horizon online policy is found using a discount factor of $\beta = 0.999$. The sum-throughput values resulting from these policies with a half-duplex relay in $N = 10$ epochs, averaged over 100 independently generated scenarios, are plotted in Figure 11. In the figure, the peak harvest rate for node T_1 , $E_{h,1}$, is varied in order to evaluate the performance of the policies at different harvesting rate scenarios. We observe that the optimal online policy, found for a horizon of $N = 10$ epochs, as well as its infinite horizon counterpart perform notably better than the naïve policies.

Finally, we compare the sum-throughput resulting from decode-and-forward, compute-and-forward, compress-and-forward, amplify-and-forward, and hybrid strategies in an EH-TWRC. The same parameters as in Figure 11 are used in simulations. The sum-throughput values obtained over a duration of $N = 10$ epochs are plotted in Figure 12. We observe that for low and high transmit powers, either decode-and-forward or compute-and-forward outperforms the other, respectively, while they both exceed the sum-throughput values of compress-and-forward and amplify-and-forward relaying. However, as expected, the hybrid strategy outperforms all single-strategy approaches, since it performs at least as good as the best one in each epoch.

VIII. CONCLUSION

In this paper, we considered the sum-throughput maximization problem in a two-way relay channel where all nodes are energy harvesting with limited battery storage, i.e., finite

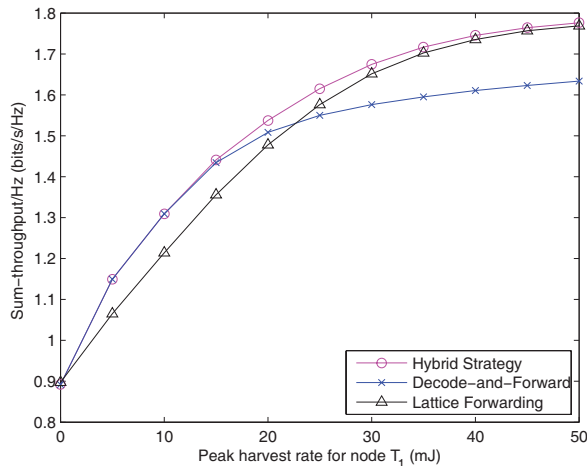


Fig. 12. Sum-throughput with various relaying strategies against peak harvest rates for node T_1 . The compress-and-forward and amplify-and-forward strategies and omitted since they perform notably worse than those in the plot.

battery. We considered decode-and-forward, compress-and-forward, compute-and-forward and amplify-and-forward relaying strategies with full-duplex and half-duplex radios. Noticing that the best relaying strategy depends on instantaneous transmit powers, we proposed a hybrid relaying scheme that switches between relaying strategies based on instantaneous transmit powers. We solved the sum-throughput maximization problem for the EH-TWRC using an iterative generalized directional water-filling algorithm. For cases where offline information about energy harvests is not available, we formulated dynamic programs which yield optimal online transmit power policies. Simulation results confirmed the benefit of the hybrid strategy over individual relaying strategies, and the improvement in sum-throughput with optimal power policies over naïve power policies. The online policies found via dynamic programming also proved to perform better than their naïve alternatives. It was observed that in a two-way channel with energy harvesting nodes, either of the communication phases, i.e., broadcast or multiple access phases, can be limiting, impacting the optimal transmit powers in the non-limiting phase as well. Thus, the jointly optimal policies were observed not to be the throughput maximizers for each individual node, or the sum-throughput maximizers for a subset of nodes - a fundamental departure in the structure of optimal policies in previous work [4]–[6], [11], [13], [14].

We remark that the offline throughput maximization problem for the full-duplex and half-duplex cases when decode-and-forward relaying is used can also be solved using the subgradient descent algorithm as shown in [41]. Future directions for this channel model include optimal offline and online power policies for more involved models with data arrivals at the sources, data buffers at the relay, or a direct channel between sources.

REFERENCES

[1] K. Tutuncuoglu, B. Varan, and A. Yener. Optimum transmission policies for energy harvesting two-way relay channels. In *ICC Workshop on*

Green Broadband Access: Energy Efficient Wireless and Wired Network Solutions, June 2013.

- [2] D. Niyato, E. Hossain, and A. Fallahi. Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization. *IEEE Trans. on Mobile Computing*, 6(2):221–236, February 2007.
- [3] J. Lei, R. D. Yates, and L. Greenstein. A generic model for optimizing single-hop transmission policy of replenishable sensors. *IEEE Trans. on Wireless Comm.*, 8(2):547–551, February 2009.
- [4] J. Yang and S. Ulukus. Optimal packet scheduling in an energy harvesting communication system. *IEEE Trans. on Comm.*, 60(1):220–230, January 2012.
- [5] K. Tutuncuoglu and A. Yener. Optimum transmission policies for battery limited energy harvesting nodes. *IEEE Trans. on Wireless Comm.*, 11(3):1180–1189, March 2012.
- [6] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener. Transmission with energy harvesting nodes in fading wireless channels: Optimal policies. *IEEE Jour. on Selected Areas in Comm.*, 29(8):1732–1743, September 2011.
- [7] C. K. Ho and R. Zhang. Optimal energy allocation for wireless communications with energy harvesting constraints. *IEEE Trans. on Signal Processing*, 60(9):4808–4818, September 2012.
- [8] B. Devillers and D. Gunduz. A general framework for the optimization of energy harvesting communication systems with battery imperfections. *Jour. of Comm. and Networks, Special Issue on Energy Harvesting in Wireless Networks*, 14(2):130–139, April 2012.
- [9] K. Tutuncuoglu, A. Yener, and S. Ulukus. Optimum policies for an energy harvesting transmitter under energy storage losses. *IEEE Jour. on Selected Areas in Comm.: Wireless Comm. Powered by Energy Harvesting and Wireless Energy Transfer*, 33(3):467–481, March 2015.
- [10] S. Luo, R. Zhang, and T. J. Lim. Optimal save-then-transmit protocol for energy harvesting wireless transmitters. *IEEE Trans. on Wireless Comm.*, 12(3):1196–1207, July 2013.
- [11] K. Tutuncuoglu and A. Yener. Communicating with energy harvesting transmitters and receivers. In *Information Theory and Applications Workshop, ITA*, February 2012.
- [12] H. Mahdavi-Doost and R. D. Yates. Energy harvesting receivers: Finite battery capacity. In *IEEE International Symposium on Information Theory, ISIT*, July 2013.
- [13] J. Yang and S. Ulukus. Optimal packet scheduling in a multiple access channel with energy harvesting transmitters. *Jour. of Comm. and Networks, Special Issue on Energy Harvesting in Wireless Networks*, 14(2):140–150, April 2012.
- [14] J. Yang, O. Ozel, and S. Ulukus. Broadcasting with an energy harvesting rechargeable transmitter. *IEEE Trans. on Wireless Comm.*, 11(2):571–583, February 2012.
- [15] O. Ozel, J. Yang, and S. Ulukus. Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery. *IEEE Trans. on Wireless Comm.*, 11(6):2193–2203, June 2012.
- [16] K. Tutuncuoglu and A. Yener. Sum-rate optimal power policies for energy harvesting transmitters in an interference channel. *Jour. of Comm. and Networks, Special Issue on Energy Harvesting in Wireless Networks*, 14(2):151–161, April 2012.
- [17] M. A. Antepi, E. Uysal-Biyikoglu, and H. Erkal. Optimal packet scheduling on an energy harvesting broadcast link. *IEEE Jour. on Selected Areas in Comm.*, 29(8):1721–1731, September 2011.
- [18] P. Blasco, D. Gunduz, and M. Dohler. Low-complexity scheduling policies for energy harvesting communication networks. In *IEEE International Symposium on Information Theory, ISIT*, July 2013.
- [19] D. Gunduz and B. Devillers. Two-hop communication with energy harvesting. In *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP*, December 2011.
- [20] O. Orhan and E. Erkip. Optimal transmission policies for energy harvesting two-hop networks. In *Conference on Information Sciences and Systems, CISS*, March 2012.
- [21] O. Orhan and E. Erkip. Energy harvesting two-hop networks: Optimal policies for the multi-energy arrival case. In *IEEE Sarnoff Symposium*, May 2012.
- [22] O. Orhan and E. Erkip. Throughput maximization for energy harvesting two-hop networks. In *IEEE International Symposium on Information Theory, ISIT*, July 2013.
- [23] I. Ahmed, A. Ikhlef, R. Schober, and R. K. Mallik. Power allocation for conventional and buffer-aided link adaptive relaying systems with energy harvesting nodes. *IEEE Trans. on Wireless Comm.*, 13(3):1182–1195, March 2014.

- [24] Y. Luo, J. Zhang, and K. B. Letaief. Optimal scheduling and power allocation for two-hop energy harvesting communication systems. *IEEE Trans. on Wireless Comm.*, 12(9):4729–4741, September 2013.
- [25] C. Huang, R. Zhang, and S. Cui. Throughput maximization for the Gaussian relay channel with energy harvesting constraints. *IEEE Jour. on Selected Areas in Comm.*, 31(8):1469–1479, August 2013.
- [26] I. Ahmed, A. Ikhlef, R. Schober, and R. K. Mallik. Joint power allocation and relay selection in energy harvesting AF relay systems. *IEEE Wireless Comm. Letters*, 2(2):239–242, April 2013.
- [27] S. J. Kim, P. Mitran, and V. Tarokh. Performance bounds for bidirectional coded cooperation protocols. *IEEE Trans. on Information Theory*, 54(11):5235–5241, November 2008.
- [28] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh. Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols. *IEEE Trans. on Information Theory*, 57(10):6405–6418, October 2011.
- [29] B. Rankov and A. Wittneben. Achievable rate regions for the two-way relay channel. In *IEEE International Symposium on Information Theory, ISIT*, July 2006.
- [30] W. Nam, S. Y. Chung, and Y. H. Lee. Capacity of the Gaussian two-way relay channel to within 1/2 bit. *IEEE Trans. on Information Theory*, 56(11):5488–5494, November 2010.
- [31] D. Gunduz, E. Tuncel, and J. Nayak. Rate regions for the separated two-way relay channel. In *Allerton Conference*, September 2008.
- [32] B. Varan, K. Tutuncuoglu, and A. Yener. Energy harvesting communications with continuous energy arrivals. In *Information Theory and Applications Workshop, ITA*, February 2014.
- [33] H. Liu, F. Sun, C. Thai, E. De Carvalho, and P. Popovski. Optimizing completion time and energy consumption in a bidirectional relay network. In *IEEE International Symposium on Wireless Communication Systems, ISWCS*, August 2012.
- [34] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, MA, 1999.
- [35] M. Gorlatova, A. Wallwater, and G. Zussman. Networking low-power energy harvesting devices: Measurements and algorithms. In *IEEE INFOCOM*, April 2011.
- [36] P. Blasco, D. Gunduz, and M. Dohler. A learning theoretic approach to energy harvesting communication system optimization. *IEEE Trans. on Wireless Comm.*, 12(4):1872–1882, April 2013.
- [37] M. B. Khuzani, H. E. Saffar, E. H. M. Alian, and P. Mitran. On optimal online power policies for energy harvesting with finite-state Markov channels. In *IEEE International Symposium on Information Theory, ISIT*, July 2013.
- [38] H. Li, N. Jaggi, and B. Sikdar. Relay scheduling for cooperative communications in sensor networks with energy harvesting. *IEEE Trans. on Wireless Comm.*, 10(9):2918–2928, September 2011.
- [39] D. P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, Belmont, MA, 1995.
- [40] M. Gorlatova, A. Bernstein, and G. Zussman. Performance evaluation of resource allocation policies for energy harvesting devices. In *International Symposium on Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks, WiOpt*, May 2011.
- [41] K. Tutuncuoglu, B. Varan, and A. Yener. Energy harvesting two-way half-duplex relay channel with decode-and-forward relaying: Optimum power policies. In *IEEE International Conference on Digital Signal Processing, DSP*, July 2013.